DOCUMENT RESUME

ED 053 552 24 EM 009 133

AUTHOR	Kolmos, Alfred S.
TITLE	Effects of Instructional Media in Teaching Beginning Statistics in a Teacher Education Program. Final Report.
INSTITUTION	Wisconsin State Universities Consortium of Research Development, Stevens Point.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C. Bureau of Research.
REPORT NO	WSU-CORD-28
BUREAU NO	BR-6-2728-28
PUB DATE	Aug 69
GRANT	OEG-3-6-062728-2129
NOTE	35p.
EDRS PRICE	EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS	Audiovisual Aids, *Autoinstructional Aids, Carrels, *Comparative Analysis, *Conventional Instruction, *Phonotape Recordings, *Slides, Tape Recordings, Teacher Education, Teaching Methods

ABSTRACT

A self-instructional program for teaching elementary descriptive statistics was developed. The program consisted of 35 mm. slide-and-tape presentations of twenty half-hour lessons. To test the effectiveness of this approach, students majoring in elementary education were divided into four treatment groups. One group used the self-instructional program on their own in a study carrel. A second group used the program in a study carrel supplemented by mandatory attendance at a 30-minute laboratory session scheduled once a week. A third treatment group used the program in the classroom and were lectured by an instructor adhering closely to the taped materials. The fourth group consisted of a traditional lecture-discussion approach to teaching. Those students in the study-carrel-only group scored significantly lower than any other group. No other significant differences were discovered between any of the other treatments. The major conclusion based on the results of this study was that self-instructional programs are not as effective as are the traditional approaches to teaching unless the self-instructional programs are supplemented by a minimum of student-instructor contact and interaction. (JY)

28

WSU-CORD

SCOPE OF INTEREST NOTICE
The ERIC Facility has assigned this document for processing to: SP EM

In our judgement, this document is also of interest to the clearinghouses noted to the right. Indexing should reflect their special points of view.

ED053552

The Wisconsin State Universities Consortium of Research Development

Research Report

EFFECTS OF INSTRUCTIONAL MEDIA IN TEACHING BEGINNING STATISTICS IN A TEACHER EDUCATION PROGRAM

Alfred S. Kolmos
Wisconsin State University - Whitewater
Whitewater, Wisconsin

Cooperative Research

Wisconsin State Universities
and the
United States Office of Education
Bureau of Research - Higher Education

Office of the Director WSU-CORD
240 Main Building
Wisconsin State University
Stevens Point, Wisconsin 54481

EM 009 133

ED053552

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

FINAL REPORT

WSU-CORD
The Consortium of Research Development
Of The
Wisconsin State Universities

Project No. 760-541-70-1007-06
Grant No. 3-6-062728-2129
Local Project No. 28

EFFECTS OF INSTRUCTIONAL MEDIA IN TEACHING BEGINNING STATISTICS IN A TEACHER EDUCATION PROGRAM

Alfred S. Kolmos
Wisconsin State University
Whitewater, Wisconsin

August 1969

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

FINAL REPORT

WSU-CORD
The Consortium of Research Development
Of The
Wisconsin State Universities

EFFECTS OF INSTRUCTIONAL MEDIA IN TEACHING BEGINNING STATISTICS IN A TEACHER EDUCATION PROGRAM

Wisconsin State University
Whitewater, Wisconsin

August 1969

The research reported herin was performed pursuant to a grant with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

TABLE OF CONTENTS

	Page
SUMMARY	iv
CHAPTER	
1. INTRODUCTION	1
2. METHODS	5
3. FINDINGS AND ANALYSIS	10
4. CONCLUSIONS AND RECOMMENDATIONS	13
REFERENCES	15
APPENDICES	
A. OBJECTIVES FOR A UNIT IN ELEMENTARY STATISTICS	17
B. CONTENT AREA FOR A UNIT IN ELEMENTARY STATISTICS	18
C. TABLE OF SPECIFICATIONS FOR A UNIT IN ELEMENTARY STATISTICS	20
D. SAMPLE SCRIPT	21
E. SAMPLE ASSIGNMENT SHEET	25
F. SAMPLE ITEMS FROM THE REVISED STATISTICS EXAMINATION	26
G. ANALYSIS OF GRADE-POINT AVERAGES DATA	28
H. ANALYSIS OF MATHEMATICS TEST DATA	29
I. ANALYSIS OF STATISTICS TEST DATA	30
TABLES	
1. RESULTS OF STATISTICS TEST	10

SUMMARY

The purpose of this study was to develop a self-instructional program for teaching elementary descriptive statistics and to test its effectiveness in attaining desired objectives. The program consisted of 35 mm slides and taped presentations of twenty half-hour lessons. Subjects were 94 junior and senior students enrolled as Elementary Education majors at Wisconsin State University - Whitewater.

One of four treatments was administered to four sections of a Measurement and Evaluation course. Treatment A consisted of instruction through the use of study carrels only; Treatment B consisted of instruction through the use of study carrels supplemented by mandatory attendance at a 30-minute laboratory session scheduled once a week; Treatment C consisted of instruction through the use of the 35 mm slides in the classroom with the instructor adhering closely to the taped material used in the carrels; Treatment D consisted of the traditional lecture-discussion approach to teaching.

Significant differences were found between Treatments A and each of the other treatments. No significant differences were discovered between any of the other treatments. The major conclusion based on the results of this study was that self-instructional programs are not as effective as are the traditional approaches to teaching unless the self-instructional programs are supplemented by a minimum of student-instructor confrontation and interaction.

CHAPTER 1

INTRODUCTION

The purpose of this study is to implement some of the concepts of an instructional systems developmental approach and to test its significance in teaching and learning. Specifically, the project is designed to determine the effectiveness of a self-instructional carrel approach in the teaching of beginning statistics, a portion of a course in Measurement and Evaluation in the Elementary School.

With the great increase in college enrollment it becomes more imperative that additional emphasis be placed on techniques and modes of instruction which prove to be more effective in nature. This is as true in teacher education as it is in other fields of endeavor. Wisconsin State University - Whitewater is a case in point. In 1960 Whitewater had 2007 students; in 1968 9402 crowded onto its premises. The number of faculty during the same period increased from 119 to 526, a slightly lower rate of increase.

Popham (1967) contends that several forces have coalesced to produce a new focus in education, a focus aimed at the preparation of validated instructional materials. These forces include the impact of programed instruction, the increased financial input into education on the part of the federal government, and growing impatience with the impact of researchers' contributions on the field of education. As a result of these forces, Popham feels that there is an increasing commitment to the systematic, empirically based development of instructional products which are demonstrably capable of bringing about specified behavior in learners.

In 1967 Schueler and Lesser (1967) emphasized the impact of instructional materials on teacher education. The purpose of their publication was to disseminate to the profession "the implications of technological developments to the future of teacher education and of their importance as potential means for the continued improvement of programs for the preparation of teachers." It was their hope that the availability of such information would stimulate and facilitate the effective study and appropriate use of the new technology. The primary question, as they saw it, was how can the functional outcomes of teacher education be furthered by the use of the instructional media?

Recently there appears to be a fresh approach to the problem of instructional media, one labeled the Instructional Systems Development. This program begins with a careful, almost microscopic, examination of expected learnings and outcomes in a particular discipline or

course. In the light of this examination, learning experiences are categorized into those that can be performed best by (1) one student working alone, or in a study carrel, (2) one, or a very few, students working closely with a professor, (3) students working in relatively small groups, with or without a professor, or in medium sized groups (20-60), and (4) large group and super large lecture-demonstrations. In essence, the instructional systems approach focuses attention upon learning and learners. It then explores resources, human and technological, which can organize learning for maximum efficiency and effectiveness. This approach also makes a careful distinction between teaching and instruction. In some instances, impersonal or mechanical media (textbooks, teaching machines, syllabi, assignment sheets, films, videotapes, etc.) can be used quite adequately for informative purposes.

It should be noted that the instructional systems approach is not a theory of instruction nor does it necessarily indicate the type of instruction to be employed. It is, rather, an integral part of any theory and can be utilized in any mode or technique of instruction deemed advisable for the accomplishment of the varied and multitudinal objectives of an instructional program. Regardless of the curriculum design employed, an instructional systems approach to planning, organizing, and implementing its procedures would seem to be advisable.

Despite Saettler's (1968) contention that comprehensive theoretical and experimental bases are lacking for a media-message design, there are indications that such a design will be given more of an emphasis in the future. Briggs et al. (1967) have made recommendations for designing a media program. Smith, Schagrin, and Poorman (1967) have applied some of the recommended principles to a pilot study concerned with the selection of instructional media in a Project Physics course with promising results.

Instructional materials are being developed in greater numbers and variety than ever before and are being designed to supplement and amplify the effects of texts, laboratories, and teacher presentations. They are also geared to facilitate a more effective use of the capacity of teachers to adjust to differences in students and in their environment. As various authors have pointed out, the role of the teacher is changing toward one as tutor and guide. The traditional function of the teacher as a purveyor of information and as a demonstrator is being transferred more and more to slides and tapes as well as to other forms of instructional media, leaving the teacher with more time to devote to a concern for individual differences and preparation for individualized instruction.

The major problems faced by the classroom teacher still remain, however. What type of learning experiences are most appropriate for the objectives desired? How can instructional media aid in accomplishing those objectives efficiently and effectively? The teacher has no specific guidelines as yet for determining what aspects of a subject should be assigned to what medium. Some attempts have been made along this line, notably by Gagné (1965) who has proposed a taxonomy

of behavioral objectives and a related taxonomy of types of learning which have implications for the selection of materials for instruction.

No one is in a better position to know what specific objectives are most appropriate for a unit of learning than is the classroom teacher. It should follow, then, that the classroom teacher is in an enviable position as far as the development of instructional materials is concerned. Developmental research is carried on with a deliberate regard to developing specific programs and demonstrating the feasibility of some instructional system or technique. Developmental research and the classroom teacher, then, are identified with identical purposes. It is hoped that this study will serve as a first step in indicating one of the avenues open to those concerned with an effective and efficient method of instruction, whether it is for the training of teachers or for any other instructional purpose. To ignore such an avenue toward increasing teacher effectiveness is to shirk one's responsibility as an educator.

In summary, this study involves the development of a self-instructional program involving elementary statistics based on the specific objectives of that unit. The study is designed to test that program in its effectiveness in accomplishing the stated objectives.

RELATED LITERATURE

Much expository material is available in current literature extolling the virtures of instructional material and a multimedia approach for effective teaching. The rationale given is difficult to deny and the arguments forwarded appear to be logical and valid. Despite this, however, there is a dearth of empirical data supporting the contentions of the instructional media proponents.

In the October 1968 issue of the Review of Educational Research, Householder (1968) notes that research in the applications of instructional media is contradictory and inconclusive with regard to the effectiveness of most media. Further, he feels that additional experimental evidence is needed to develop guidelines which indicate the types of educational purposes best served by each of the instructional media. It is quite surprising to note the scant research evidence that Householder, as well as others, presents in his summary.

Shemick (1965) utilized 35 mm slides and tapes in the instruction of metal spinning. He found that the subjects who received their instruction by means of this technique required less aid from the instructor but they required more time, took more trials, and did not attain the performance quality attained by those students who participated in the more conventional approach. Moeller (1967), in using 8 mm film and tape recordings, came up with similar results in teaching elementary engine lathe operation. In contrast to those findings,

however, Training in Business and Industry (1966) found that the time required to train industrial assembly workers was reduced 50 per cent when the training utilized color slides and tape recordings.

Davis (1961), at the elementary school level, described procedures she followed in teaching 5th and 6th grade Science and 1st through 3rd grade Reading. She used primarily the tape recorder in her self-instructional device but offered no experimental evidence as to its effectiveness. The same comment is true of Young (1961) who used tapes and slides in the presentation of life experienced by American Field Service representatives.

Torkelson and Driscoll (1968) in their summary of the utiliization and management of learning resources have little to offer in supporting or refuting the value of instructional media. They cite studies concerning the assessment of programed instruction, language laboratories, motion pictures, and educational television but none are particularly relevant to a multimedia approach.

Schueler and Lesser (1967) discovered in their investigation of media research that the "no significant difference" was quite common. Their summary, although focusing on the role of media in teacher education, centered around the use of television, programed instruction, and language laboratories. No research was cited which might apply to the problem outlined in this research.

CHAPTER 2

METHODS

A portion of a course in Measurement and Evaluation in the Elementary School was selected for this study, specifically that portion devoted to elementary descriptive statistics. A period of ten 50-minute sessions is usually devoted to this block of instruction in the traditional lecture-discussion type of presentation. The selection of this topic for the study was chosen deliberately for one major reason. In teaching statistics to undergraduate students in this context, there are those students who enter the unit with a trepidation and apprehension which borders on a mental block. There are those, also, who because of their mathematical prowess and/or past experiences are able to move rapidly through the unit. Still others, of course, are somewhere between these extremes. Faced with this problem, the instructor tends to set a median pace, a pace which does not meet specifically the needs of those at either end of the distribution.

It was envisioned that the application of auto-instruction would free the more capable student to move through the unit smoothly and quickly. This approach would also free the other students to procede at their own pace and remain with a concept until they had that concept mastered to their own satisfaction. This would enable them, also, to seek out the instructor in areas of difficulty. They would not move on to the next step until they were ready. It was envisioned, also, that auto-instruction would free the instructor so he could work more closely and carefully with that section of the class which needed his attention and guidance.

Before conducting the experiment itself it was necessary to define in behavioral terms the objectives for this portion of the course, to delineate specifically the content area involved, and to prepare a table of specifications to serve as a guide in the preparation of a test to measure the effectiveness of the various treatments in the experiment. Bloom (1956), Gronlund (1966), and Noll and Saupe (1965) were used as references to aid the preparation of those documents. The documents are included in Appendices A, B, and C.

Instructional media in the form of tape recordings, 35 mm slides, and assignment sheets were developed for student use in six study carrels. The tape recordings were accomplished by the investigator in order to eliminate the variable of instructor personality as much as possible. Further, tape recordings were accomplished for

twenty 25-minute sessions rather than ten 50-minute sessions. It was felt that such an organization would promote a greater and more efficient utilization of the carrels. There very well could be occasions where a student would have an opportunity to accomplish a 25-minute lesson but not sufficient time available for a longer unit of instruction. A sample of a transcript of one of the recordings is included in Appendix D. A total of 310 35 mm slides were developed for the program and six copies of each slide were reproduced to insure the availability of any lesson desired by the student. Assignment sheets were developed for ten of the lessons. The assignment sheets were designed to give students limited practice in some of the techniques employed in elementary statistics courses of this nature. A sample of one of the assignment sheets in included in Appendix E.

Lien's (1967) Measurement and Evaluation of Learning and Noll's (1965) Introduction to Educational Measurement were required texts for the course. In addition, three hand-outs were prepared and distributed to supplement the treatment of certain topics presented in the two texts.

Finally, an objective examination was constructed based, of course, on the objectives previously determined. The 60-item multiple choice examination was first critiqued by Dr. Norman E. Gronlund, Professor of Educational Psychology, University of Illinois, and then tested on 129 Elementary Education students at Wisconsin State University - Whitewater before being revised and edited for utilization in this project. Sample items from the revised examination are included in Appendix F.

The subjects for the experiment were 94 junior and senior students enrolled at Wisconsin State University - Whitewater. All of the students majored in Elementary Education and were in one of four sections of the Measurement and Evaluation course. The individual sections were selected randomly to participate in one of four treatments.

Treatment A consisted of instruction through the use of the study carrels only. Students were advised that they could expect no assistance from the instructor nor could they have more time available for the viewing of any lesson than the 25-minute period allotted. These procedures were deemed necessary to insure that achievement through Treatment A was primarily a function of the program per se.

Treatment B consisted of instruction through the use of the study carrels supplemented by mandatory attendance at a 25-minute laboratory session scheduled once a week. The student achievement under Treatment B would thus be a function of the self-instructional program plus the additional time devoted to the subject and the effect of the instructor variable during the laboratory sessions. The short laboratory session was designed specifically to solve

student problems or misconceptions and to answer any questions concerning the course material. No attempt was made to review or summarize specific aspects of the course.

Treatment C consisted of scheduled classes with the instructor employing the same 35 mm slides, assignment sheets, and hand-outs utilized in the study carrels. Student achievement under this treatment would be a result of the visual media employed in addition to the instructor variable. The instructor made a concerted effort to adhere as closely as possible to the taped presentation utilized in the carrels. This was done to negate the effect of a different content or approach to the subject matter.

Treatment D consisted of the traditional approach to college teaching with the instructor employing primarily blackboard and chalk as instructional media. No attempt was made to adhere to the taped presentation. This group was considered the control group for this experiment. Achievement for this group was considered a direct result of the instructor and his presentation.

It should be noted that the variable of time as a factor was inconsequential except for Treatment B where an additional 25-minute laboratory session was scheduled. The two classes experiencing Treatments C and D met for three 50-minute sessions a week for a total of ten meetings. This time was identical to the time allotted for Treatment A, instruction through the use of the study carrels only.

HYPOTHESES

The general hypothesis for this study is that students utilizing study carrels supplemented by mandatory attendance at scheduled laboratory sessions will achieve desired objectives as well as those students attending traditional lecture-discussion classes. This hypothesis is generated under the assumption that elementary statistics is a field of study wherein informing or the acquisition of data is of prime importance. It is felt that teaching for insightful questioning, illumination, and evaluation, while important, is not as vital in the study of very elementary statistics as it is in other portions of the course in Measurement and Evaluation.

In order to test the general hypothesis and in order to test the value of instructional media, the following hypotheses were postulated:

(1) There will be a difference in the mean scores of students under Treatments A and B. Testing of this hypothesis will help to determine whether or not students participating in a self-instructional program will benefit from supplementing such instruction with seminar sessions. The major variable in this comparison is one of time.

Students under Treatment B will have had additional time devoted to the same material. The effect of the instructor variable at the seminar sessions is deemed to be of secondary importance.

(2) There will be a difference in the mean scores of students under Treatments A and C. Testing of this hypothesis will help to determine the effectiveness of taped programs in the imparting of instruction. The major variable in this comparison is that of instructor personality. Students under both treatments will have been taught desired concepts through the use of the 35 mm slides.

(3) There will be a difference in the mean scores of students under Treatments A and D. Testing of this hypothesis will aid in determining the effectiveness of a carrel approach to teaching in comparison with the more conventional or traditional approach.

(4) There will be a difference in the mean scores of students under Treatments B and C. Testing of this hypothesis will help to determine the effectiveness of a carrel program supplemented by scheduled seminars in comparison with a program wherein visual media is used extensively in the classroom. The major variables operating in this comparison are those of time and instructor personality.

(5) There will be a difference in the mean scores of students under Treatments B and D. Testing of this hypothesis will help to determine the effectiveness of the modified carrel program in comparison with the more conventional classroom approach. The variables in this instance consist of time devoted to the material and the mode of presentation.

(6) There will be a difference in the mean scores of students under Treatments C and D. Testing of this hypothesis will help to determine the effectiveness of the use of slides in a traditional teaching program. The variables compared in this instance are those of the prepared slides and script versus the traditional mode of presentation.

STATISTICAL TREATMENT

Grade-point averages of the students were compared to test for the homogeneity of the students in the four sections of the Measurement and Evaluation course. Grade-point averages of the students in the four sections and an analysis of variance of the GPA data is included in Appendix G. No significant difference was found in the analysis.

It was envisioned that students' scores on the mathematics portion of the American College Testing Program (ACT) would be used to compare the four groups for homogeneity of mathematical ability. Scores for all students were not available, however, so a test

developed by Dr. Robert E. Stake, Professor Educational Psychology, University of Illinois, was used in its stead. Scores of the students in the four sections and an analysis of variance of the mathematics data is included in Appendix H. No significant difference was found in the analysis.

Raw scores on the statistics examinations were used to determine whether there was a significance in the means of the scores of the students in the four groups. An analysis of variance was utilized for this purpose and, as for the GPA and mathematics data, a 5 per cent level was arbitrarily selected as being significant. A discussion of this analysis is included in the next chapter.

CHAPTER 3

FINDINGS AND ANALYSIS

The general hypothesis for this study was that students utilizing study carrels supplemented by mandatory attendance at scheduled laboratory sessions would achieve desired objectives as well as those students attending the traditional lecture-discussion classes. In order to test the general hypothesis and in order to test the value of instructional media, six hypotheses were generated and four treatments were administered, one to each of the four sections in Measurement and Evaluation in the Elementary School.

The investigator constructed a 60-item objective test in elementary statistics and the resulting scores were used to support or reject the null hypotheses. Table 1 summarizes the results for each of the four treatments.

TABLE 1

RESULTS OF STATISTICS TEST

SECTION	N	MEAN	SD
A	23	35.70	4.5914
B	26	40.15	4.1015
C	23	39.13	7.0668
D	22	39.32	6.1155

An analysis of variance of the data indicates that a significant difference exists at the .05 level and, therefore, that a null hypothesis must be rejected. In order to determine which of the differences between these means are significant and which are not, the Duncan new multiple range test was utilized. The analysis of variance for this data and the application of the Duncan new multiple range test may be found in Appendix I.

The results will be discussed in the order in which they were stated in Chapter 2. The first null hypothesis was that there would be no difference in the mean scores of students under Treatments A and B. The null in this case must be rejected. Students in both Treatments A and B participated in the use of the carrels as their primary mode of instruction. The only difference in the

treatments was that in Treatment B the students met once a week for a thirty-minute period and had an opportunity to question the instructor and to clear up any misconceptions they may have had. On the basis of this experiment, then, it may be concluded that additional time spent on the same subject material will aid students in achieving higher grades in a testing situation. The results may also lead to the conclusion that in a self-instructional program student-instructor confrontation and interaction, even at a minimal level, will raise student test grades significantly. This, it is assumed, will prove of some comfort to those who may fear that automation will remove some teachers from their traditional instructional role.

The second hypothesis, stated in null terms, was that there would be no difference in achievement between those students taught through Treatments A and C. Treatment A consisted of instruction through the use of the study carrels only; Treatment C consisted of instruction through the use of the 35 mm slides utilized in the carrels but the teaching was conducted in the conventional classroom locale with the investigator adhering closely to the script employed in the self-instructional program. Here, again, the null hypothesis had to be rejected. This would appear to substantiate one of the conclusions in the preceding paragraph, student-instructor confrontation and interaction will raise test grades significantly. Inasmuch as the same period of time was involved in both treatments and inasmuch as the same materials were presented in a very similar approach, this seems to be the only conclusion one can reach. The only other possible conclusion concerns the self-instructional program itself but such a discussion is more appropriate in a later paragraph.

The third hypothesis generated was that there would be no difference in mean scores between students undergoing Treatments A and D. Stated in different terminology, there would be no difference in achievement between those students participating in the carrel program only and those in the traditional lecture-discussion situation. Results indicated that, once more, the null hypothesis has to be rejected. There was a difference in the mean scores of students undergoing the two treatments. The results here would further substantiate the conclusions reached in the previous paragraphs. In this study, at least, a carrel approach to teaching is less effective than the more conventional or traditional approach. The results in this instance may cast some doubt on the quality of the carrel program per se but, again, consideration of this aspect is relegated to a later paragraph.

Results of the achievement test indicated that the remaining null hypotheses could not be rejected. In the first instance, comparisons were made between the results obtained under Treatments B and C. Under Treatment B the carrel program was supplemented by mandatory attendance at a 30-minute laboratory session scheduled

once a week. The instructor used the 35 mm slides in Treatment C. The fact that no significant difference existed between these two treatments leads credence to a contention that achievement gained through additional time devoted to subject matter (in the case of Treatment B) can be compensated for by insuring more student-instructor confrontation and interaction. Both treatments were basically the same as far as content and approach were concerned, so no conclusions can be drawn as far as the quality of the self-instructional program is concerned.

The rejection of the null hypotheses in the final two sets of treatments, however, leads one to conclusions regarding the merits of the carrel program per se. The fact that under Treatment B (the modified carrel approach) students achieved as much as those undergoing Treatment D (the traditional classroom lecture-discussion approach) leads one to logically infer that the carrel program materials in this experiment were at least as good as the materials used in the lecture-discussion program. The addition of time under Treatment B leading to higher achievement is apparently balanced by the greater amount of student-instructor interaction under Treatment D.

In the final instance, the fact that the null hypothesis could not be rejected means that the addition of slides to a lecture-discussion situation did not effect any significant gain in achievement in this experiment. Under Treatment C the investigator not only used the 35 mm slides used in the carrels, but he also adhered closely to the script used in the self-instructional laboratory. A no significant difference between Treatment C and the control treatment results in the conclusion, again, that the materials used in the carrel program were at least as good as those used in the lecture-discussion treatment. Neither the amount of time nor the amount of instructor-student interaction was relevant under these treatments.

CHAPTER 4

CONCLUSIONS AND RECOMMENDATIONS

Based on the results of this study the following conclusions can be substantiated:

(1) Additional time devoted to the study of subject material will aid in achieving objectives.

(2) The greater the amount of time spent in student-instructor confrontation and interaction, the more likely that students will attain desired objectives.

(3) Less time devoted to the study of subject material can be compensated for by insuring a greater amount of time spent in student-instructor interaction.

(4) Achievement restricted to a self-instructional carrel approach is not as effective as the traditional lecture-discussion approach to teaching and learning.

(5) The mere utilization of 35 mm slides in a traditional classroom setting does not guarantee increased achievement.

(6) It appears possible that a classroom teacher, following pre-determined objectives and content, can construct a self-instructional program which will result in as much student achievement as his more traditional classroom approach. This conclusion would seem to be especially justified if the course objectives were primarily concerned with the dissemination of data. Additional research along these lines, however, is needed before complete acceptance of this conclusion can be expected.

Several implications are warranted from the results of this investigation. The advantages of self-instructional programs are not only self-evident but have been expounded upon by their proponents. The fact that a student may progress at his own speed, the fact that concepts are learned before more complex concepts are initiated, the fact that self-instructional programs provide versatile and flexible accomodations are quite well accepted. In regard to this last advantage, it might be pointed out that in this investigation several members of the sections undergoing the carrel approach were absent for nearly a week while on a University sponsored music tour. Because of the carrel approach, these students were able to accomplish course requirements with a minimum of discomforture. It thus behooves

a dedicated educator to utilize such programs to his advantage, knowing that student achievement will not suffer as a result.

In a teacher training institution, especially, students should not only be aware of varied instructional techniques, but should experience them as well. It is reasonable to expect that students who have profitable and enjoyable experiences in self-instructional programs in college will tend to initiate and utilize such techniques when they become responsible for instruction. This hypothesis, in fact, could well be the basis for further study.

The value to a teacher who constructs a self-instructional program should be obvious. Before attempting such a project, the teacher must be well informed as to his objectives and he will more critically evaluate course content, textbooks, and references in the light of those objectives. He will spend considerable time in organizing his presentations and in allowing for varied types of instructional media and he will construct better, more valid and more reliable, instruments for measurement purposes. Even if the teacher's program were not put into operation, it would be safe to assume that his qualities as a teacher would improve measurably. Although measurement of good teaching is an extremely difficult task, the foregoing remarks might also serve as a basis for future investigations.

Finally, more effort needs to be expended by instructional media personnel in order to arrive at a more satisfactory theory regarding their products. A satisfactory, comprehensive theory should guide researchers along lines which would be most fruitful. It is not enough to laud the merits of a multi-media approach to learning. Users must be confident that such an avenue is not only theoretically sound but is substantiated by empirical research.

REFERENCES

Bloom, Benjamin S. (ed.), Taxonomy of Educational Objectives, (New York: David McKay Company, Inc., 1956).

Briggs, Leslie J. et al., Instructional Media: A Procedure for the Design of Multi-Media Instruction, A Critical Review of Research, and Suggestions for Future Research, Monograph No. 2 (Pittsburgh, Pa.: American Institutes for Research, 1967).

Davis, Nancy H., "A Second Teacher." Audiovisual Instruction, Vol. 6, No. 4, April 1961.

Gagné, Robert M., The Conditions of Learning, (New York: Holt, Rinehart and Winston, Inc., 1965).

Gronlund, Norman E., Instructor's Manual for Measurement and Evaluation in Teaching, (New York: The Macmillan Company, 1966).

Householder, Daniel L., "Techniques and Modes of Instruction." Review of Educational Research, October 1968.

Lien, Arnold J., Measurement and Evaluation of Learning, (Dubuque, Iowa: Wm. C. Brown Company, 1967).

Moeller, Carl A., "A Comparison of Selected Audio-Visual Methods and Lecture Demonstration Methods in Teaching Manipulative Skills Related to Metal Working Operations." Journal of Industrial Teacher Education, March 1967.

Popham, W. James, "Instructional Product Development: Two Approaches to Training." AV Communication Review, XV (Winter 1967).

Noll, Victor H., Introduction to Educational Measurement (2d ed.), (Boston: Houghton Mifflin Company, 1965).

Noll, Victor H. and Saupe, Joe L., Instructor's Manual to Accompany Introduction to Educational Measurement, Second Edition, (Boston: Houghton Mifflin Company, 1965).

Saettler, Paul, "Design and Selection Factors." Review of Educational Research, April 1968.

Schueler, Herbert and Lesser, Gerald S., Teacher Education and the New Media, (Washington, D. C.: American Association of Colleges for Teacher Education, 1967).

Shemick, John M., "Teaching a Skill by Machine." Industrial Arts and Vocational Education, October 1965.

Smith, Daniel M., Schagrin, Morton, and Poorman, Eugene L., "Multi-media Systems: A Review and Report of a Pilot Project." AV Communication Review, XV (Winter 1967).

Torkelson, Gerald M. and Driscoll, John P., "Utilization and Management of Learning Resources." Review of Educational Research, April 1968.

Training in Business and Industry, "Kodak Evolves Better Assembly Training." Training in Business and Industry, August 1966.

Young, Kenneth G., "Strengthening International Friendships through Tapes and Slides." Audiovisual Instruction, Vol. 6, No. 8, October 1961.

APPENDIX A

OBJECTIVES FOR A UNIT IN ELEMENTARY STATISTICS

1. Knows statistical terms and facts

    1.1 Defines terms and facts
    1.2 Recalls properties of statistical terms
    1.3 Distinguishes between statistical terms
    1.4 Recognizes appropriate formulae

2. Understands statistical terms and concepts

    2.1 Describes appropriate utilization of terms and concepts
    2.2 Solves simple statistical problems
    2.3 Constructs appropriate tables and graphs
    2.4 Translates graphical data into verbal terms
    2.5 Recognizes inappropriate utilization of data
    2.6 Identifies relationships between concepts

3. Applies knowledge of facts and concepts

    3.1 Solves simple statistical problems

        3.11 Indicates appropriate statistic to use
        3.12 Selects correct formula
        3.13 Computes correct answer using formula

    3.2 Provides rationale for solving more complex problems

4. Interprets concepts from novel situations

    4.1 Provides prognosis from given situation

APPENDIX B

CONTENT AREA FOR A UNIT IN ELEMENTARY STATISTICS

A. Introduction/Terminology (½ hour)

1. The Origin of Statistics
2. Meaning of Statistics
3. Terminology
4. Misuses of Statistics

B. Frequency Distributions (½ hour)

1. Array
2. Simple Ranking
3. Grouping

C. Graphic Representation (1 hour)

1. The Histogram
2. The Frequency Polygon
3. Ogives

D. Measures of Central Tendency (2½ hours)

1. The Mean
2. The Median
3. The Mode
4. Characteristics of the Mean, Median, and Mode

E. Measures of Position (½ hour)

1. Percentiles
2. Related Measures (Deciles, Quartiles)

F. Measures of Variability (2 hours)

1. The Range
2. Mean Deviation
3. The Quartile Deviation (Semi-Interquartile Range)
4. The Standard Deviation
5. The Variance
6. Use of Variability Measures

G. Measures of Relationship - Correlation (2 hours)

1. The Product-Moment Coefficient of Correlation
2. Correlation from Ranks

H. Standard Scores (1 hour)

1. z-scores
2. T-scores
3. IQ scores
4. Stanines

APPENDIX C

TABLE OF SPECIFICATIONS FOR A UNIT IN ELEMENTARY STATISTICS

CONTENT \ OBJECTIVES	1 Knows statistical terms and facts	2 Understands statistical terms and concepts	3 Applies knowledge of facts and concepts	4 Interprets concepts from novel situations	TOTALS	
					Test Items	Percentage of Items
A. Introduction/Terminology	2*	1	-	-	3	5
B. Frequency Distributions	2	1	-	-	3	5
C. Graphic Representation	3	2	1	-	6	10
D. Measures of Central Tendency	5	4	3	3	15	25
E. Measures of Position	1	1	$\frac{1}{2}$	$\frac{1}{2}$	3	5
F. Measures of Variability	4	4	2	2	12	20
G. Measures of Relationship	5	3	2	2	12	20
H. Standard Scores	2	2	$\frac{1}{2}$	$1\frac{1}{2}$	6	10
TOTALS — Test Items	24	18	9	9	60	
TOTALS — Percentage of Items	40	30	15	15		100

* Numbers refer to the number of test items in the sixty-item test.

APPENDIX D

SAMPLE SCRIPT

Elementary Statistics
Lesson 2: Frequency Distributions

2-1*This is the second of twenty half-hour lessons in elementary statistics, a portion of the course Measurement and Evaluation in the Elementary School.

You have recently completed the second test in this course, we have gone over the examination in class, and you have discovered how many of the test items you had completed correctly. This number, the number of items you answered correctly, is your raw score. A raw score, by itself, is comparatively meaningless. It was to emphasize this point that I refused to inform you of the scores that other members of your section received. What was your raw score?2-2 35? 42? 48? Are you satisfied with that score? How well did you do in comparison with the rest of the class? Would you expect your score to be equivalent to an A? A B? A C? By attempting to answer these questions, you should arrive at a conclusion that realistic answers are impossible uless some descriptions of the scores of the entire class is given. This description of scores is what we're concerned with in descriptive statistics. It's that simple!

In order to understand more readily what is to follow, let me give you a concrete example.2-3 This slide represents the scores made by a previous class in Measurement and Evaluation. I listed the scores here just as I have them in my Class Record Book. Looking at these scores quickly, it is difficult to determine whether a score of 42 is one to be satisfied with or not. By presenting all of the scores of the class, I have given you, however, a description of the scores, but it's a pretty inefficient one.

It would be much more efficient if I presented the scores to you, not in a random order, but in an order from high scores to low scores. This is what I've done in the next slide,2-4 and this is what is called an array. Now you can determine quite readily how satisfied you might be with a raw score of 42; 13 students out of a class of 50 have higher scores than you. That's not too bad, is it? Some of you'd be real happy with such a score, I know.

By arranging the scores in an array we can better evaluate the worth of a particular score. We can further evaluate any score by noting its rank order. This next slide illustrates what I'm talking

---

*Superscripts refer to the 35 mm slide number and are inserted in the script where that slide should be displayed.

about.[2-5] Obviously, the student who achieved the raw score of 50 ranks first in the class and the student who obtained 49 ranks second. If there were no duplication of scores, this ranking procedure would be an extremely simple one, but unfortunately there are duplications of scores. There were three students who had 45 as their score. Since it would not be correct to say that one ranks higher than the others, it is necessary to average the three ranks, 3, 4, and 5, and assign the average, 4, to all three scores. In like manner, students who obtained a raw score of 44 are assigned a rank order of 6.5, the average of ranks 6 and 7. As rank orders are computed for the remainder of the students, it is determined that the student who obtained the raw score of 30, the lowest score, has a rank order of 50, for there were a total of 50 students in the class. Assigning rank order to students according to their scores, then, enables us to gain a much better concept of the worth of any particular raw score.

When working with many scores, arranging those scores into an array can be a tedious and time-consuming job. One way out of this difficulty is to arrange the scores into what is called a frequency distribution. Such a distribution consists of two columns, as illustrated in the next slide.[2-6] The first column indicates the various scores in the distribution in order of size. The second column indicates the number of times each score occurs, or its frequency. It might be pointed out at this time that we usually indicate frequency by the small letter "f", and the total number of scores by the capital letter "N", in this case 50. In this frequency distribution, then, we have described all of the scores obtained by the students, and we have described them a little more efficiently than we had before.

We're still not out of the woods, though. Many times the range of scores is so great that an additional process is required. There sometimes is so wide a range of scores that it becomes more economical to group the scores according to size. For instance, and this is a realistic example, suppose that raw scores from a particular test range from a low of 44 to a high of 112.[2-7] Were we to list all of the scores in an array we would have quite a long list. By grouping the scores into class intervals of five, for instance, we could save ourselves considerable effort, for we'd end up with fewer scores, or rather, groups of scores. In our example, we could group the scores so that the highest class interval includes all the scores from 110 through 114; the next class interval would include scores from 105 through 109, and so on. The complete grouping arrangement is usually referred to as a grouped frequency distribution.

There are several steps to follow in making the grouped frequency distribution.[2-8] We can use our original set of scores to illustrate the process. (1) First, determine the range of scores. The lowest raw score was 30; subtracting that score from the highest score, 50, we obtain a range of 20.[2-9] (2) The second step is to divide the range by the number of class intervals or groups we desire. There is no absolutely fixed rule for the number of class intervals; different authors advise different numbers. Some recommend from 10 to 20, some

12 to 15, Noll suggests from 8 to 16. I haven't seen any figure less than 8 nor any figure larger than 20, so we'll just leave it at that. In any case, it is well to remember that the fewer number of class intervals, the easier it is to make computations, but the less accurate we are. The more class intervals we use, the more accurate we are, but the more work that is involved. Going back to the example, then, let's decide on 8 as the number of class intervals we want. This will make it slightly easier for instructional purposes than it would be if we used a greater number of groups. Dividing the range, 20, by the number of groups we want, 8, we find the size of each of the class intervals is 2.5. Since it is impractical, usually, to use any class interval except a whole number, the class interval to use in our example could be either 2 or 3. As we'll see later, odd-numbered intervals are usually preferred over even-numbered intervals. This is because with odd-numbered intervals, the mid-point of the class interval will be a whole number. The mid-point of the class interval, 90-94, for instance, is 92, a whole number. With even-numbered intervals, however, the mid-point will be a fraction. The mid-point of the class interval, 92-95, for instance, is 93.5. Computations later will be less complex if we deal with whole numbers. As a result, then, we'll choose 3 as the size of our class interval.

2-10(3) The next step is to make a table of class intervals which will include all scores. To facilitate the tabulation, it is usually customary to start each class with a multiple of the class interval. Our lowest score in the example is 30, the size of the class interval is 3, so we have no problem at all, for 30, of course, is a multiple of 3. Had the lowest score been 31, we would have still used 30 to start the lowest class, for 31 is not a multiple of 3. The lowest class interval will be 30-32. Note that it does include three possible scores, 31, 31, and 32, so the size of the class interval is indeed 3. The next higher class interval will be 33-35, to include scores of 33, 34, and 35. We continue the table of class intervals until we find the class interval which includes the highest obtained score. The highest score we've already seen to be 50, so the highest class interval will be 48-50.

Let me digress for a moment in order to point out one other property of class intervals with which we'll be concerned later, that is, the limits of the class intervals. The lowest class interval in our example is 30-32, the next higher is 33-35. Theoretically, I suppose, a student could obtain a raw score of 32.8. Into which class interval would we put this score?2-11 In statistics it is conventional to define the class limit as midway between the highest value indicated in the lower class interval and the lowest value indicated in the next higher class interval. In other words, the limits of the lowest class interval will be 29.5 and 32.5, or more precisely, 29.5 and 32.4999+; the limits of the next higher class interval will be 32.5 and 35.5, and so on.

Now, let us get back to our grouped distribution again.2-12 The next step (4) is to tally the scores one by one into the proper

class interval, and finally the last step (5) is to add the tallies and write the sums opposite each interval under a column labeled "f" or "frequencies." As a check, we can add up the frequencies to make sure that we've not missed any tallies along the way. We started with 50 scores, and we'd better still have 50 scores. As you can see, we do.

Let me review for you the steps in making a grouped frequency distribution[2-13]:

(1) Determine the range. Merely subtract the lowest score from the highest.

(2) Divide the range by the number of class intervals desired (somewhere between 8 and 20).

(3) Make a table of class intervals which will include all scores. Start each class with a multiple of the class interval.

(4) Tally the scores into the proper class intervals.

(5) Add the tallies in each class interval and write the sums under a frequency column. Check your results by adding all of the frequencies to help insure you've made no error.

This concludes the lesson. If you're a little confused over what has been presented,[2-14] check Noll, pages 38-40 and 457-462 and Lien, pages 111-113. Before starting Lesson 3, obtain the Assignment Sheet for Lesson 2 and follow the procedures strictly. The closer you adhere to the given instructions, the easier it will be for you to review your work and to complete further assignments.[2-15]

APPENDIX E

SAMPLE ASSIGNMENT SHEET

ASSIGNMENT SHEET LESSON 2

Ed. Found. 424 Elementary School Measurement and Evaluation

The following lists represent the test grades made by four sections of the Measurement and Evaluation course.

1. Arrange each section's grades in an array. In other words, you will have four different arrays.

2. Tabulate the grades in each of the sections into four different frequency distributions. Indicate each score and the number of times it occurred.

3. Tabulate the grades in each of the sections into four grouped frequency distributions using an interval of 3. This will result in fewer class intervals than we ordinarily desire, but for illustrative purposes it will facilitate computations. In Section II the lowest class interval will be 45-47, in Section III the lowest class interval will be 33-35, in Section V the lowest class interval will be 36-38, and in Section VI the lowest class interval will be 42-44.

4. Before beginning Lesson 3, check your work with the instructor at the self-instructional lab.

5. Retain this assignment sheet, for succeeding assignments will be based on the data presented.

Section II (N=29)		Section III (N=27)		Section V (N=30)		Section VI (N=22)	
45	54	48	52	49	54	49	49
54	47	53	52	51	50	54	48
49	54	53	51	43	48	49	43
48	51	50	52	51	53	51	50
45	49	50	53	44	46	50	55
54	51	47	43	45	50	52	59
55	51	49	49	56	47	53	52
49	48	48	52	39	55	51	
48	48	53	52	47	44	54	
50	48	43	34	49	56	48	
51	48	50	50	43	52	54	
52	51	51	48	45	49	56	
53	51	44		38	52	57	
51	56	53		45	49	51	
55		54		54	45	46	

APPENDIX F

SAMPLE ITEMS FROM THE REVISED STATISTICS EXAMINATION

Indicate the item on the following list which is an example of an ordinal scale.

(1) Fahrenheit temperatures
(2) loudness
(3) occupations
(4) popularity

The main reason for grouping data in class intervals is to

(1) mask the identity of the persons involved.
(2) permit the caluculation of statistical measures in addition to the mean.
(3) reduce the frequency of clerical errors.
(4) reduce the amount of labor involved.

An ogive is most useful in determining

(1) standard scores.
(2) percentiles.
(3) standard deviations.
(4) correlation coefficients.

For the two questions below, there are listed occasions for the use of the various measures of central tendency. Indicate "1" on your answer sheet if the mean is the most appropriate measure of central tendency, "2" if the median is most appropriate, and "3" if the mode is most appropriate.

It is desired that the size of every measure in the distribution be taken into consideration.

The distribution involves values which have a nominal scale only.

A teacher administered an intelligence test to a group of fourth graders and obtained the following IQs: 88, 115, 110, 105, 90, 140, 115. What was the median IQ for these seven pupils?

(1) 105
(2) 109
(3) 110
(4) 115

A teacher graded a teacher-made test administered to her 15 pupils. She found that the raw scores ranged from 30 to 90. If she assumed a normal distribution of scores, she could logically conclude that the standard deviation of the scores would be approximately

(1) 3
(2) 4
(3) 6
(4) 10

The mean score on a standardized test has been determined to be 50 with a standard deviation of 6. This means that approximately 34 per cent of those taking the test will obtain scores ranging from

(1) 50 to 62.
(2) 47 to 53.
(3) 44 to 56.
(4) 44 to 50.

Which of the following coefficients indicates the highest degree of correlation?

(1) -.90
(2) -.50
(3) 0.00
(4) +.30

The manual for the high school battery of the Iowa Test in Education Development states there is a correlation of .62 between the composite scores on the ITED and the scholarship averages in the first year of college. If Mary scored high on the ITED, this would appear to indicate that she

(1) definitely will do well in college.
(2) has a 62 per cent chance of succeeding in college.
(3) will do well in college because she did well on the ITED.
(4) is quite likely to do well in college.

The formula $\frac{x}{\sigma}$ can be used to calculate

(1) percentiles.
(2) T-scores.
(3) z-scores.
(4) stanines.

Which one of the following scores represents the highest level of achievement?

(1) T-score = 59
(2) z-score = 1.0
(3) Stanine = 6
(4) Percentile rank = 86

APPENDIX G

ANALYSIS OF GRADE-POINT AVERAGES DATA

	N	$\bar{X}$	SD	ΣX	ΣX^2
Section A	23	2.5843	.3741	59.44	156.8332
Section B	26	2.8658	.3383	74.51	216.5039
Section C	23	2.6691	.3399	61.39	166.5155
Section D	22	2.7282	.4498	60.02	168.1974

ANALYSIS OF VARIANCE

Source of Variation	df	Sum of Squares	Mean Square	F
"Between" Groups	3	1.0356	.3452	2.3356 (P > .05)
"Within" Groups	90	13.3045	.1478	
Total	93	14.3401		

APPENDIX H

ANALYSIS OF MATHEMATICS TEST DATA

	N	$\bar{X}$	SD	ΣX	ΣX^2
Section A	23	11.13	4.6747	256	3352
Section B	26	10.65	4.1505	277	3399
Section C	23	10.48	3.9768	241	2889
Section D	22	11.68	4.6259	257	3473

ANALYSIS OF VARIANCE

Source of Variation	df	Sum of Squares	Mean Square	
"Between" Groups	3	19.8992	6.6331	0.3344 (F>.05)
"Within" Groups	90	1785.0051	19.8334	
Total	93	1804.9043		

APPENDIX I

ANALYSIS OF STATISTICS TEST DATA

ANALYSIS OF VARIANCE

Source of Variation	df	Sum of Squares	Mean Square	F
"Between" Groups	3	274.5772	91.5257	2.8467 ($P < .05$)
"Within" Groups	90	2893.6356	32.1515	
Total	93	3168.2128		

DUNCAN'S NEW MULTIPLE RANGE TEST ANALYSIS
(Entries indicate range factors)

SECTION	MEAN	A	C	D	B	APPROPRIATE SIGNIFICANT RANGE FACTORS
A	35.70		16.464*	17.158*	21.983*	$R_2 = 15.876$
C	39.13			0.901	5.039	$R_3 = 16.709$
D	39.32				4.050	$R_4 = 17.265$
B	40.15					

*$p < .05$, df = 90, k = 2, 3, 4